

Peer-Review

Huang, Henry. 2026. "Navigating Parameter Space: Mitigating Catastrophic Forgetting in Continual Learning." *Journal of High School Science* 10 (3): 331–57. <https://doi.org/10.64336/001c.165645>.

1. The manuscript presents Range Matrix as a new parameter-regularization paradigm in which every parameter is constrained within an allowable range rather than only those deemed important by the Fisher Information Matrix. However, this conceptual framing appears closely related to existing interval-based continual learning methods, particularly the Hyperrectangle Training framework of Wołczyk et al. (ICML 2022), which likewise represents feasible regions of parameter space through interval constraints applied to all network parameters. Although the mathematical implementation differs—Range Matrix derives parameter bounds from local loss linearization whereas Hyperrectangle Training employs interval arithmetic and nested hyperrectangles—the underlying scientific principle of constraining the entire parameter space rather than selectively regularizing important weights appears substantially similar. Consequently, the manuscript should explicitly compare Range Matrix with this prior work and clearly articulate what is genuinely novel about its formulation. Without such discussion, the current novelty claims are difficult to evaluate and may overstate the originality of the proposed approach.

2. Going back to point 1, conceptually your idea resembles several other existing ideas:

- trust-region optimization-
- projected gradient methods
- proximal optimization
- gradient clipping
- box-constrained optimization
- interval parameter regularization
- parameter projection

The manuscript never demonstrates that Range Matrix is fundamentally different from these methods, and therefore, novel.

3. The algorithm should be evaluated against at least several representative continual-learning methods, for example

- SI (Synaptic Intelligence)
- MAS
- RWalk
- LwF
- replay

Otherwise it is impossible to determine whether Range Matrix advances the state of the art or merely performs similarly to one early baseline.

4. The manuscript should report multiple seeds with mean accuracy, standard deviation and confidence intervals. Without variance estimates, it is impossible to know whether the reported 4% improvement is statistically meaningful.

5. The manuscript assumes the clipping imposed by the Range Matrix is beneficial, but it does not investigate how the method behaves when the new task requires large parameter changes because it is substantially different from previous tasks. In such situations, the optimizer may repeatedly clip updates, limiting the network's capacity to acquire the new knowledge. Consequently, the algorithm may be inherently better suited to gradual or closely related continual-learning scenarios than to highly dissimilar tasks. Evaluating performance as task similarity decreases would therefore be an important experiment to determine whether Range Matrix disproportionately sacrifices plasticity in exchange for memory stability.

This could be investigated by progressively increasing the dissimilarity or amount of new information presented in each continual-learning step. For example, compare performance when introducing:

- one new class,
- two new classes,

- five new classes,
- an entirely unrelated dataset.

The resulting degradation in new-task accuracy would quantify the algorithm's effective learning capacity. This metric should be compared against established continual-learning methods such as EWC, SI, MAS, replay-based methods, and interval-constrained approaches to determine whether Range Matrix sacrifices plasticity more rapidly than existing techniques.

6. Does the feasible parameter space shrink over time until continual learning eventually stalls?

Since the algorithm continually intersects or tightens allowable ranges after each task, the admissible region of parameter space may become progressively smaller. This could impose a theoretical limit on the number or diversity of tasks the model can learn. Studying how the "volume" of the feasible parameter space evolves across tasks would be a valuable extension of this work and could reveal an inherent scalability limitation of the proposed method.

This could be evaluated by training on a long sequence of tasks (far beyond five CIFAR-10 contexts) while monitoring:

- new-task accuracy,
- forgetting,
- average allowable parameter range,
- percentage of parameters at their clipping boundaries,
- total feasible parameter-space volume (or an approximation thereof).

If the allowable region asymptotically approaches zero, the algorithm may possess an intrinsic capacity limit independent of network size. This would be a very important theoretical property.

7. The paper currently evaluates accuracy, but Range Matrix is really a geometric algorithm: it defines and updates a feasible region in parameter space. Accuracy is only an indirect consequence of that geometry. A stronger paper would therefore measure the evolution of the geometry itself. For example:

- average interval width,
- fraction of clipped updates,
- feasible parameter-space volume,
- distance from parameter boundaries,
- rate at which the feasible region contracts.

Those measurements would reveal why the algorithm succeeds or fails, not merely whether it achieves a particular accuracy.

I encourage the authors to include these analyses because they transform the paper from "we propose another continual-learning algorithm" into "we characterize the dynamics of constrained parameter spaces in continual learning." That would substantially strengthen the scientific contribution.

8. How much of the performance improvement is due to the regularization algorithm, and how much is simply due to giving the network more parameters?

To isolate the effect of the Range Matrix algorithm, the experiments should ideally answer questions such as:

- If both architectures had the same number of parameters, would Dynamic FC still outperform Static FC?
- How much of the accuracy gain comes from architectural expansion versus the regularization method itself? Without controlling for that, architecture and algorithm are intertwined.

The manuscript does not disentangle these effects. A stronger experimental design would either:

- compare methods at equal parameter counts, or
- explicitly quantify how much additional capacity the Dynamic FC introduces and discuss its contribution to performance. This is an important confounding factor that deserves discussion.

9. P appears to function as a global control over the network's plasticity. As P increases, the allowable parameter ranges expand, enabling greater adaptation to new tasks but increasing forgetting. Conversely, decreasing P tightens the allowable parameter ranges, improving memory retention at the expense of learning capacity. This suggests that P is not merely a hyperparameter

but a quantitative regulator of the stability–plasticity tradeoff. The manuscript would be strengthened by explicitly characterizing P in this way and experimentally quantifying the relationship between P , new-task learning, memory retention, and long-term continual learning performance. This would be a much more scientifically interesting interpretation of the algorithm than simply presenting P as another hyperparameter to tune.

First, I would like to thank the reviewers for their time and insightful suggestions. Overall, this manuscript’s purpose is to investigate the effect of the Static vs. Dynamic FC layer choice on model performance. The contribution of the Range Matrix is, in the context of this manuscript, to investigate the behaviors noticed during experimentation with FC layers. A future work will seek to investigate the Range Matrix method itself closer, analyzing concerns such as stability vs. plasticity as well as comparing it against other continual learning approaches.

Reviewer #1

1. “The manuscript presents Range Matrix as a new parameter-regularization paradigm in which every parameter is constrained within an allowable range rather than only those deemed important by the Fisher Information Matrix. However, this conceptual framing appears closely related to existing interval-based continual learning methods, particularly the Hyperrectangle Training framework of Wołczyk et al. (ICML 2022), which likewise represents feasible regions of parameter space through interval constraints applied to all network parameters. Although the mathematical implementation differs—Range Matrix derives parameter bounds from local loss linearization whereas Hyperrectangle Training employs interval arithmetic and nested hyperrectangles—the underlying scientific principle of constraining the entire parameter space rather than selectively regularizing important weights appears substantially similar. Consequently, the manuscript should explicitly compare Range Matrix with this prior work and clearly articulate what is genuinely novel about its formulation. Without such discussion, the current novelty claims are difficult to evaluate and may overstate the originality of the proposed approach.”

Response:

The reviewer points out a key observation for parameter regularization methods, in that they are all designed to constrain parameter space. The main difference between methods lies within the way these constraints are calculated. As suggested by the reviewer, the Range Matrix differs from the Hyperrectangle Training framework in that while the hyperrectangle work approximates parameter space as a rigid geometric object (in this case a hyperrectangle), Range Matrix calculates constraints individually for each parameter using loss gradient. Additionally, the authors of Hyperrectangle Training framework note a key limitation in that their approach is not compatible with batch normalization, which is common in neural networks. The Range Matrix possesses no such limitation. The Range Matrix was primarily developed to investigate limitations found in EWC Static FC for class incremental learning. This comparison is included in the revised prior work section under section 2.2 “Continual Learning approaches” in the third paragraph.

2. “Going back to point 1, conceptually your idea resembles several other existing ideas:

- trust-region optimization-
- projected gradient methods
- proximal optimization
- gradient clipping
- box-constrained optimization
- interval parameter regularization
- parameter projection

The manuscript never demonstrates that Range Matrix is fundamentally different from these methods, and therefore, novel.”

Response:

Thanks to the reviewer for pointing out these approaches. The ideas presented were researched thoroughly, and an effort was made to find continual learning papers utilizing these approaches. These methods either approach the issue by modeling the parameter space as a geometric object, utilize EWC-like fisher importance values, or utilize static clipping values. Range Matrix differs in that it utilized loss gradients to calculate permissible ranges for each parameter at the end of each context. A discussion about these papers was added in the prior work section 2.2 “Continual Learning approaches” in the third paragraph.

3. “The algorithm should be evaluated against at least several representative continual-learning methods, for example

-SI (Synaptic Intelligence)

-MAS

-RWalk

-LwF

-replay

Otherwise it is impossible to determine whether Range Matrix advances the state of the art or merely performs similarly to one early baseline.”

Response:

As mentioned in the response to the reviewer’s first question, the purpose of this paper was to identify the role of the FC layer in model performance in class incremental learning. Range Matrix was developed to investigate the limitations found through that study. At the current stage, this manuscript is not trying to claim that Range Matrix improves the state of the art. I agree with the reviewer that the Range Matrix needs to be compared to other continual learning methods to fully reveal its potential. The limitation of how Range Matrix was not evaluated against other continual learning methods was actually addressed in the Limitations and Future Directions section (5) of the original paper. This limitation was expressed more clearly in the revised manuscript in Section 5. The contributions of the paper were also made clear at the end of the introduction (section 1). Finally, the conclusion (section 6) was also revised to summarize the paper’s findings more clearly.

4. The manuscript should report multiple seeds with mean accuracy, standard deviation and confidence intervals. Without variance estimates, it is impossible to know whether the reported 4% improvement is statistically meaningful.

Response:

This was a good suggestion by the reviewer, and the Range Matrix and EWC were rerun with 5 additional seeds. Both Range Matrix and Static FC were utilizing Static FC layers. The results and confidence intervals are reported in the revised Evaluation Results section 4.3.

5. “The manuscript assumes the clipping imposed by the Range Matrix is beneficial, but it does not investigate how the method behaves when the new task requires large parameter changes because it is substantially different from previous tasks. In such situations, the optimizer may repeatedly clip updates, limiting the network’s capacity to acquire the new knowledge. Consequently, the algorithm may be inherently better suited to gradual or closely related continual-learning scenarios than to highly dissimilar tasks. Evaluating performance as task similarity decreases would therefore be an important experiment to determine whether Range Matrix disproportionately sacrifices plasticity in exchange for memory stability.

This could be investigated by progressively increasing the dissimilarity or amount of new information presented in each continual-learning step. For example, compare performance when introducing:

-one new class,

-two new classes,

-five new classes,

-an entirely unrelated dataset.

The resulting degradation in new-task accuracy would quantify the algorithm's effective learning capacity. This metric should be compared against established continual-learning methods such as EWC, SI, MAS, replay-based methods, and interval-constrained approaches to determine whether Range Matrix sacrifices plasticity more rapidly than existing techniques.

Does the feasible parameter space shrink over time until continual learning eventually stalls? Since the algorithm continually intersects or tightens allowable ranges after each task, the admissible region of parameter space may become progressively smaller. This could impose a theoretical limit on the number or diversity of tasks the model can learn. Studying how the "volume" of the feasible parameter space evolves across tasks would be a valuable extension of this work and could reveal an inherent scalability limitation of the proposed method.

This could be evaluated by training on a long sequence of tasks (far beyond five CIFAR-10 contexts) while monitoring:

- new-task accuracy,
- forgetting,
- average allowable parameter range,
- percentage of parameters at their clipping boundaries,
- total feasible parameter-space volume (or an approximation thereof).

If the allowable region asymptotically approaches zero, the algorithm may possess an intrinsic capacity limit independent of network size. This would be a very important theoretical property. The paper currently evaluates accuracy, but Range Matrix is really a geometric algorithm: it defines and updates a feasible region in parameter space. Accuracy is only an indirect consequence of that geometry. A stronger paper would therefore measure the evolution of the geometry itself. For example:

- average interval width,
- fraction of clipped updates,
- feasible parameter-space volume,
- distance from parameter boundaries,
- rate at which the feasible region contracts.

Those measurements would reveal why the algorithm succeeds or fails, not merely whether it achieves a particular accuracy.

I encourage the authors to include these analyses because they transform the paper from "we propose another continual-learning algorithm" into "we characterize the dynamics of constrained parameter spaces in continual learning." That would substantially strengthen the scientific contribution.

How much of the performance improvement is due to the regularization algorithm, and how much is simply due to giving the network more parameters?

To isolate the effect of the Range Matrix algorithm, the experiments should ideally answer questions such as:

-If both architectures had the same number of parameters, would Dynamic FC still outperform Static FC?

-How much of the accuracy gain comes from architectural expansion versus the regularization method itself? Without controlling for that, architecture and algorithm are intertwined.

The manuscript does not disentangle these effects. A stronger experimental design would either:

- compare methods at equal parameter counts, or
- explicitly quantify how much additional capacity the Dynamic FC introduces and discuss its contribution to performance. This is an important confounding factor that deserves discussion.

P appears to function as a global control over the network's plasticity. As P increases, the allowable parameter ranges expand, enabling greater adaptation to new tasks but increasing forgetting.

Conversely, decreasing P tightens the allowable parameter ranges, improving memory retention at the expense of learning capacity. This suggests that P is not merely a hyperparameter but a quantitative regulator of the stability-plasticity tradeoff. The manuscript would be strengthened by explicitly characterizing P in this way and experimentally quantifying the relationship between P,

new-task learning, memory retention, and long-term continual learning performance. This would be a much more scientifically interesting interpretation of the algorithm than simply presenting P as another hyperparameter to tune.”

Response:

Thank you to the reviewer for the detailed suggestions. As stated before, the main purpose of this manuscript was to explore the effect of the FC layer on performance. These experiments revealed that EWC with Static FC had 0 memory capability. From there, Range Matrix was introduced to investigate these behaviors. Thus, investigating the Range Matrix further is not within the scope of this paper. While the Range Matrix may not be ready to be investigated deeper as of now, I agree with the reviewer in that these studies are needed to reveal the true capabilities and limitations of Range Matrix. A discussion about the stability/plasticity tradeoff was added in section 3.4. Similarly to other CL methods which impose bounds on parameter space, the stability/plasticity tradeoff is important. Because the Range Matrix continues to “shrink” the parameter space as new contexts are learned, it is only logical to assume that eventually the space will become too tiny for meaningful learning. In a future work, the learning plasticity of Range Matrix will be studied through collecting metrics such as size, rate of contraction, percentage of parameters that needed to be clipped, etc. It will also investigate learning two functionally different tasks, through either more contexts or entirely different datasets, to understand the memory stability of Range Matrix. These items are all discussed in the Limitations and Future Works (section 5).

Regarding the question about FC layer vs. regularization algorithm, this manuscript includes an experiment where EWC was implemented with a Resnet model. The architecture and all else were held constant, with only the FC layers being changed, with one model running with Static and the other with Dynamic. The results showed that Dynamic FC improved over Static FC. Interestingly, utilizing a different regularization algorithm (Range Matrix) did not affect the overall performance of Dynamic FC and Static FC architectures too much. This could suggest that the architecture may be more important than regularization algorithm, at least for these two parameter regularization methods in class incremental learning. The results of this finding were made clearer in the Evaluation Results (section 4.3).

Regarding the role of P, I agree with the reviewer that P is a control over plasticity. This work simply ran a hyperparameter search of P to demonstrate that Range Matrix could have memory capability over EWC. I agree that future investigation of P’s role in the stability vs. plasticity tradeoff is needed. This future investigation is suggested in the revised manuscript in the Limitations and Future Works (section 5). The suggested characterization of P was also applied to the Evaluation Results section (section 4.3) and Range Matrix introduction section (section 3.4).

Reviewer #2

“Please verify that all links to the references point to the correct source.”

Response:

All links in the references were verified to point to the correct source.

“The authors must disclose and acknowledge any assistance received in the preparation of this manuscript, including but not limited to editorial, technical, analytical, or writing support. All such contributions must be clearly stated in the Acknowledgments section.”

Response:

An acknowledgements section (7) was added after the Conclusion section (6). Any assistance received in the preparation of the manuscript were detailed in the section.

“Include enough recent references along with foundational ones. Ensure references directly support your claims. Avoid “padding.” Use credible sources (peer-reviewed journals, reputable books, official reports).”

Response:

The prior work section, specifically section 2.2, was revised with sources. For Architecture based and Replay methods, which were not the focus of this paper, 3 extra sources were added for each, covering more recent references along with well-established methods. For the third paragraph on section 2.2 which was focused on parameter regularization, 5 sources were added to put the concept of the paper in context with similar methods.

“All assumptions (including implicit assumptions) must be explicitly stated and clearly justified. The authors should explain why each assumption is reasonable and discuss its impact on the results and conclusions.”

Response:

The manuscript was thoroughly reviewed, and assumptions were clarified and justified as necessary. The main purpose of this manuscript was to investigate Static vs. Dynamic FC. Through these experiments, EWC with Static FC was identified to have 0 memory capability. Thus, Range Matrix was developed and the results showed that it had memory capability. This is the extent to the paper’s claims on Range Matrix’s superiority, as further comparison against other continual learning methods are required to reveal its performance across the field. Furthermore, discussion was made on the plasticity vs. stability tradeoff. Because the Range Matrix conceptually reduces the parameter space each new context learned, it is logical to assume that eventually plasticity could decrease to the point where continual learning ceases. This discussion is made in the Range Matrix introduction section (section 3.4). This plasticity will be investigated through studying features such as range size, rate of parameter space contraction, percentage of parameters clipped etc. Stability can be investigated through learning different tasks through more contexts or entirely different datasets. These items are discussed in the Limitations and Future Works section (section 5).

“Avoid overreaching conclusions that extend beyond what is supported by the data and analysis.”

Response:

The overall contribution of the paper is to explore the effect of the Static/Dynamic FC architecture on model performance. The proposed Range Matrix method was developed specifically to improve upon the limitations of EWC Static FC in class incremental learning. The point was not to assert that Range Matrix improves the state of the art over other continual learning methods. The limitation of only comparing Range Matrix to EWC was addressed in the Limitations and Future directions in section 5. [reiterate future work]. The Introduction (section 1), Limitations and Future directions (section 5), as well as Conclusion (section 6) were all revised to reflect these message more clearly.

“Verify that you have used past perfect tense and third person throughout the manuscript wherever applicable.”

Response:

The manuscript was reviewed thoroughly and past perfect tense and third person were verified to be used wherever they were applicable.

Thank you for addressing my comments. The manuscript is much improved. However, there is still considerable scope for further improvement. Please see below:

While the manuscript demonstrates catastrophic forgetting qualitatively through accuracy-distribution plots and reports overall accuracy, these measures alone do not fully characterize continual-learning performance or permit rigorous comparison with the broader literature. The claim that Range Matrix preserves memory better than EWC would be substantially strengthened by reporting established continual-learning metrics, including average forgetting (loss in performance on previously learned tasks after subsequent training), backward transfer (BWT) to

quantify whether learning new tasks harms or benefits prior tasks, forward transfer (FWT) to assess whether prior learning facilitates acquisition of new tasks, and a retention ratio (final accuracy divided by peak accuracy for each previously learned task). In addition, reporting new-task accuracy (plasticity) alongside these memory metrics would explicitly quantify the stability-plasticity tradeoff. Such metrics would allow readers to distinguish improvements arising from enhanced memory retention versus improved acquisition of new tasks and would enable direct comparison with established continual-learning methods reported in the literature. Finally, because the manuscript concludes that Range Matrix preserves memory better than EWC, it would be valuable to quantify forgetting directly rather than relying primarily on qualitative accuracy-distribution figures. These additions would considerably strengthen both the rigor of the conclusions and the manuscript’s relevance to the continual-learning community.

The manuscript should provide a complete description of the training protocol to ensure reproducibility. In its current form, several essential implementation details are either omitted or insufficiently described. A machine learning study should explicitly report the number of training epochs per context, optimizer, learning rate, batch size, weight decay (or other regularization parameters), learning-rate scheduler (if used), and the stopping criterion (e.g., fixed epochs, early stopping, or convergence threshold). These details are necessary for readers to reproduce the reported results and to determine whether the observed performance differences arise from the proposed methodology rather than undocumented training choices. If these parameters are only available in the implementation code or are omitted entirely, the reproducibility and scientific rigor of the work are substantially diminished.

The comparison between Static FC and Dynamic FC simultaneously changes two important properties of the network: (i) the number of trainable parameters increases as new output neurons are added, and (ii) task representations become partitioned into largely context-specific neurons. Consequently, the current experiments cannot determine whether the observed improvement is primarily attributable to increased model capacity or to reduced interference arising from task-specific representations. An informative control experiment would be to construct a Dynamic FC architecture in which new neurons are added as before, thereby maintaining the same parameter count, but the representation of each new context is intentionally distributed across both existing and newly added neurons rather than being confined primarily to the new outputs. This experiment would isolate the effect of architectural expansion from that of context partitioning. Such an analysis would substantially strengthen the conclusion that the FC architecture itself, rather than increased capacity alone, is responsible for the observed reduction in catastrophic forgetting.

One possible explanation for the observation that Range Matrix provides little improvement over EWC when using a Dynamic FC architecture is that Dynamic FC inherently partitions the classifier into task-specific neurons. Under these conditions, the Fisher importance estimates used by EWC and the parameter constraints imposed by Range Matrix may operate on largely the same subset of parameters, leading to similar effective weight adjustments. This hypothesis could be tested by quantifying the overlap between Fisher-important parameters and Range Matrix-constrained parameters, or by analyzing the distribution of clipped updates across layers and contexts. Such an analysis would provide mechanistic insight into why the two methods converge in performance under Dynamic FC.

Firstly, thank you to the reviewer for their time and extremely detailed suggestions. While addressing the comments of the reviewer, a bug was noted in the experimentation implementation. This bug occurred in the Dynamic FC architecture and caused the FC layer neurons to not be trained. The bug was only limited to the FC layer neurons in Dynamic FC. After being resolved, all affected experiments were re-run. Overall, the accuracy of methods using

Dynamic FC saw an increase of a few percentage points. Most importantly however, the overall claims and behaviors noticed in the paper did not change as a result of this bug being fixed. The figures were all updated to reflect these new results. This was also used as an opportunity to conduct a visual pass over some of the figures, improving their readability and clarity. Overall, the Results section (Section 4) saw the most update, with it being split into multiple subsections for clarity. The results of the Hyperparameter searches are first reported in section 4.3. Then, the accuracy distributions, overall accuracy graphs, and accuracy comparison between Range Matrix and EWC static FC are discussed in section 4.4. Finally, in section 4.5, the continual-learning metrics are reported.

A new section (A study of FC Architecture) was added after the Results (Section 4). This section is designated to explore the FC Layer architecture (Specifically Dynamic FC) in greater detail, following the suggestions of the reviewer in the third and fourth comments.

Besides addressing the comments of the reviewer, the paper as a whole was passed over for minor editorial revisions for improved tone and clarity.

Reviewer #1:

“While the manuscript demonstrates catastrophic forgetting qualitatively through accuracy-distribution plots and reports overall accuracy, these measures alone do not fully characterize continual-learning performance or permit rigorous comparison with the broader literature. The claim that Range Matrix preserves memory better than EWC would be substantially strengthened by reporting established continual-learning metrics, including average forgetting (loss in performance on previously learned tasks after subsequent training), backward transfer (BWT) to quantify whether learning new tasks harms or benefits prior tasks, forward transfer (FWT) to assess whether prior learning facilitates acquisition of new tasks, and a retention ratio (final accuracy divided by peak accuracy for each previously learned task). In addition, reporting new-task accuracy (plasticity) alongside these memory metrics would explicitly quantify the stability-plasticity tradeoff. Such metrics would allow readers to distinguish improvements arising from enhanced memory retention versus improved acquisition of new tasks and would enable direct comparison with established continual-learning methods reported in the literature. Finally, because the manuscript concludes that Range Matrix preserves memory better than EWC, it would be valuable to quantify forgetting directly rather than relying primarily on qualitative accuracy-distribution figures. These additions would considerably strengthen both the rigor of the conclusions and the manuscript’s relevance to the continual-learning community.”

Thank you to the reviewer for this comment. As suggested, Range Matrix and EWC with both FC architectures were evaluated with BWT, Average Forgetting, Retention Ratio, and New-Task Accuracy (plasticity). Regarding the Forward Transfer metric, standard FWT measures zero-shot performance on future tasks to determine if learning prior tasks facilitates quicker acquisition of new knowledge. However, because this paper strictly focuses on the Class Incremental Learning setting, future tasks introduce entirely different classes. For the Dynamic FC setup, the model physically lacks the output neurons to classify these unseen contexts prior to training, meaning zero-shot accuracy is effectively 0%. Additionally, Class Incremental Learning literature traditionally focuses on forgetting of previously learned tasks, and thus commonly omits Forward Transfer as a metric. Overall, the metrics of BWT, Average Forgetting, and Retention ratio serve to directly address the reviewer’s suggestion of representing the catastrophic forgetting quantitatively, while New-Task Accuracy is reported to provide insights into plasticity. The Results section (section 4) was reorganized into separate subsections, with these plasticity/stability metrics introduced in Section 4.2 and reported in section 4.5. The results verify that Range Matrix improves memory capability for Static FC architectures, with Range Matrix Static FC showing results such as an average forgetting of 33% to EWC Static FC’s nearly 96% average forgetting score.

“The manuscript should provide a complete description of the training protocol to ensure reproducibility. In its current form, several essential implementation details are either omitted or insufficiently described. A machine learning study should explicitly report the number of training epochs per context, optimizer, learning rate, batch size, weight decay (or other regularization

parameters), learning-rate scheduler (if used), and the stopping criterion (e.g., fixed epochs, early stopping, or convergence threshold). These details are necessary for readers to reproduce the reported results and to determine whether the observed performance differences arise from the proposed methodology rather than undocumented training choices. If these parameters are only available in the implementation code or are omitted entirely, the reproducibility and scientific rigor of the work are substantially diminished.”

Response:

Initial Learning rate, training epochs, batch size, stopping criterion, Optimizer (and optimizer settings), and Learning Rate scheduler were included in Appendix 5. These settings were reported for both Range Matrix and EWC, and the settings for each continual learning method remained constant regardless of the FC architecture.

The comparison between Static FC and Dynamic FC simultaneously changes two important properties of the network: (i) the number of trainable parameters increases as new output neurons are added, and (ii) task representations become partitioned into largely context-specific neurons. Consequently, the current experiments cannot determine whether the observed improvement is primarily attributable to increased model capacity or to reduced interference arising from task-specific representations. An informative control experiment would be to construct a Dynamic FC architecture in which new neurons are added as before, thereby maintaining the same parameter count, but the representation of each new context is intentionally distributed across both existing and newly added neurons rather than being confined primarily to the new outputs. This experiment would isolate the effect of architectural expansion from that of context partitioning. Such an analysis would substantially strengthen the conclusion that the FC architecture itself, rather than increased capacity alone, is responsible for the observed reduction in catastrophic forgetting.

Response:

Thank you to the reviewer for this insightful comment. Following the suggestion, a study was run where both EWC and Range Matrix were trained with a Dynamic FC architecture that expanded after each context but did not possess any form of context partitioning (the entire FC layer was trained, instead of only the added neurons). The experiment methodology and results are described in section 5.1. The results of the experiment showed that when context partitioning was removed, the performance of both continual learning methods suffered heavily from catastrophic forgetting. Specifically, EWC with no context partitioning experienced average forgetting of 95.51%, while Range Matrix experienced average forgetting of 94.45%. This demonstrates that it is the FC architecture design itself, not the increased model capacity, that accounts for the reduction in catastrophic forgetting.

One possible explanation for the observation that Range Matrix provides little improvement over EWC when using a Dynamic FC architecture is that Dynamic FC inherently partitions the classifier into task-specific neurons. Under these conditions, the Fisher importance estimates used by EWC and the parameter constraints imposed by Range Matrix may operate on largely the same subset of parameters, leading to similar effective weight adjustments. This hypothesis could be tested by quantifying the overlap between Fisher-important parameters and Range Matrix-constrained parameters, or by analyzing the distribution of clipped updates across layers and contexts. Such an analysis would provide mechanistic insight into why the two methods converge in performance under Dynamic FC.

Response:

Once again, thank you to the reviewer for the suggestion. The first suggestion, quantifying the overlap between Fisher-important parameters and Range Matrix constrained parameters, was chosen. After each context, the Fisher Information Matrix and Range Matrix were saved. These two importance estimates were compared using Top-K overlap, computing the overlap between the Top K highest Fisher Information Matrix values and the Top K most restrictive (smallest) Range Matrix ranges. The methodology and results of this experiment are reported in Section 5.2. The results show that even the Top 25% most important parameters in both methods only share around a 49% overlap after the last context, suggesting that the definition of “importance” between the Range

Matrix and EWC methods still differ. This could mean the reason why both methods converge in performance under Dynamic FC is because the FC architecture itself is the dominant factor in deciding performance.

Thank you for comprehensively addressing my comments. Accepted.